For universals (but not finite-state learning) visit the zoo

For universals (but not finite-state learning) visit the zoo

Geoffrey K. Pullum and Barbara C. Scholz

School of Philosophy, Psychology and Language Sciences,
University of Edinburgh, Edinburgh, EH8 9AD, Scotland, United Kingdom.
gpullum@ling.ed.ac.uk; bcscholz@gmail.com

Abstract: Evans & Levinson's (E&L's) major point is that human languages are intriguingly diverse rather than (like animal communication systems) uniform within the species. This does not establish a "myth" about language universals, or advance the ill-framed pseudo-debate over universal grammar. The target article does, however, repeat a troublesome myth about Fitch and Hauser's (2004) work on pattern learning in cotton-top tamarins.

The take-home message from the target article by Evans & Levinson (E&L) is not in their title but in their subtitle. They don't show that language universals are a myth; their point is that what makes human languages really interesting for cognitive science is their diversity, not their uniformity. Boas would have endorsed this view, but it seems fresh and novel in the current context. You want species-wide universal grammar? Visit the zoo. Study puttynose monkeys (Arnold & Zuberbühler 2006), or your cat. What human beings bring to animal communication is not rigid universals but a flexible ability to employ any of a gigantic range of strikingly varied systems. That seems to be what E&L are saying.

Regrettably, though, the authors repeat a wildly false claim about results on syntactic learning in nonhuman primates. E&L were apparently misled by a statement in the literature they critique: they cite Fitch and Hauser (2004) as having demonstrated that cotton-top tamarins have "impressive learning powers over FSGs [finite state grammars]" (sect. 6, para. 3).

This meme is spreading, alarmingly: E&L even cite a paper by a brain researcher (Friederici 2004) who is looking for distinct neural systems for processing FSGs on the one hand and PSGs (phrase-structure grammar) on the other. The truth is that no one has shown monkeys to have any general ability to learn finite-state (FS) languages. It is extremely unlikely that anyone ever will. FS parsing is powerful; it would suffice for pretty much all the linguistic processing that humans ever do.

Reflect for a moment on the likelihood that tamarins could be habituated to strings matching this regular expression:

$$a\ (c^*\ d\ c^*\ d\ c^*)^*\ a + b\ (c^*\ d\ c^*\ d\ c^*)^*\ b$$

All such strings begin with either *a* or *b*; the middle is an indefinitely long stretch of *c* and *d* in random order, but always containing an even number of *d*; and strings end with whatever they began with (notice, an unbounded dependency!).

This language has a very simple FSG, but passively learning it from being exposed to examples (*acca*, *accdcccddccccda*; *bdccccdb*; *acdccddcccddda* . . .) is surely not plausible. Figuring out the grammar would surely be way beyond the abilities of any mammal other than a skilled human puzzle-solver with pencil and paper.

People have unfortunately been confusing FS languages with a vastly smaller proper subset known as the strictly local (SL) languages (Pullum & Rogers 2006; Pullum & Scholz 2007; Rogers & Pullum 2007). Fitch and Hauser unwittingly encouraged the error by remarking that FSGs "can be fully specified by transition probabilities between a finite number of 'states' (e.g., corresponding to words or calls)" (Fitch & Hauser 2004, p. 377). The equation of states with words here is an error. States in FSGs are much more abstract. Languages that can be described purely in terms of transitions between particular terminal symbols (words or calls or whatever) are SL languages. (It is not clear why Fitch and Hauser mentioned the orthogonal issue of transition probability.)

The SL class is the infinite union of a hierarchy of SL_k languages ($k > 1$), the k setting the maximal distance over which dependencies can be stated. The most basic SL languages are the SL_2 languages, describable by finite sets of *bigrams*. The SL_2 class is right at the bottom of several infinite hierarchies within the FS languages (see Pullum & Rogers [2006] or Rogers & Pullum [2007] for the mathematics).

Fitch and Hauser found that cotton-top tamarins could be habituated to an SL_2 pattern, namely the one denoted by (ab)*. They remark that perhaps tamarins fail to learn non-FS PSG languages "because their ability to differentiate successive items is limited to runs of two" (Fitch & Hauser 2004, p. 379), conceding the point that a limitation to recognizing bigrams might well be involved. Their results do not in any way imply that monkeys can acquire arbitrary FS languages from exposure to primary data. They may not be much better at pattern learning than your cat. E&L have unfortunately contributed to the spread of a myth.

As for the supposed myth of E&L's title, that of language universals, we see little prospect of sensible debate at this stage. People trying to set one up usually depict a clash between Chomsky, who has purportedly "shown that there is really only one human language" (Smith 1999 p. 1), and Joos, who is alleged to have claimed that languages may "differ from each other without limit and in unpredictable ways" (e.g., Smith 1999, p. 105).

But Chomsky (in Kasher 1991, p. 26) says merely that if all parameters of syntactic variation are "reducible to lexical properties" and if we ignore all parameters that are so reducible (hence we ignore all parameters), there is no syntactic variation at all, so the number of distinct syntactic systems is 1. This is not an empirical claim about human languages; it is a tautology. And Joos (1966, p. 96), while setting a phonology paper in historical context, merely alluded to an "American (Boas) tradition" that valued cataloguing language features over explanatory speculation. The passage quoted does not endorse that tradition or extend it to syntax.

It should be obvious that we must assume languages may differ in unpredictable ways: we do not know the limits of variation, so fieldwork often brings surprises. That was Boas's point. But equally obviously, not all conceivable differences between languages will be attested. Logically there could be dekatransitive verbs (taking ten obligatory object NPs), but there are not, because using them would outstrip our cognitive resources. In that sense there will be all sorts of limits.

This does not look like the seeds of an interesting debate, so it is just as well that E&L do not really try to pursue one. Their conclusions are not about universally quantified linguistic generalizations being mythical, but about how "the diversity of language is, from a biological point of view, its most remarkable property" (sect. 8, para. 6, their thesis 1). That is an interesting thought, and it deserves extended consideration by linguistic and cognitive scientists.

References

Arnold, K. & Zuberbühler, K. (2006) Language evolution: Semantic combinations in primate calls. *Nature* 441(7091):303.

Fitch, W. T. & Hauser, M. (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303(5656):377–80.

Friederici, A. (2004) Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Science* 8:245–47.

Joos, M., ed. (1966) *Readings in linguistics I*, 2nd edition. University of Chicago Press.

Kasher, A. (1991) *The Chomskyan turn*. Blackwell.

Pullum, G. K. & Rogers, J. (2006) Animal pattern learning experiments: Some mathematical background. Unpublished ms, University of Edinburgh. Available at: http://ling.ed.ac.uk/~gpullum/MonkeyMath.pdf

Pullum, G. K. & Scholz, B. C. (2007) Systematicity and natural language syntax. *Croatian Journal of Philosophy* 7(21):375–402.

Rogers, J. & Pullum, G. K. (2007) Aural pattern recognition experiments and the subregular hierarchy. Paper presented at the Mathematics of Language 10 Conference, UCLA, July 2007. Available at: http://ling.ed.ac.uk/~gpullum/MoL10paper.pdf

Smith, N. (1999) *Chomsky: Ideas and ideals*. Cambridge University Press.

ACKNOWLEDGMENT

We thank Rob Truswell for helpful critical comments on an earlier draft. He bears no responsibility for the opinions expressed in this one, or for any errors.